DPにおけるある解法について(数理システムにおける最適化理論とその応用)

黒岩, 大史

数理解析研究所講究録
第 899 巻 1995 年 47-52

DP におけるある解法について

新潟大学大学院自然科学研究科　黒岩 大史*(DAISHI KUROIWA)

Abstract

通常の多段決定過程では，ある時刻における状態と決定により，次の時刻の状態が（確率的に）１つ決まる．このときの最適値関数は，次の状態に対する期待値を考え，目標は，最適値関数を最適にするような政策を見つけることである．

本報告においては，ある時刻における状態と決定により与えられる許容推移状態集合の中から，次の時刻の状態が何らかの法則で選ばれるような多段決定過程を考える．その法則で代表的なものは，決定者を最も有利にするように選ばれるものと，決定者を最も不利にするように選ばれるものの２つであろう．その例として，前者は，決定者が状態を選べるときであり，後者は，ゲームの問題になっているときが考えられる．本報告で観察するのは，前者の場合であり，その最適値関数と，最適値関数を最適にするような政策について考察する．

1　問題の定式化

まず，次のような N 段の Dynamic Programming を考える：

S_n：状態空間で，Banach 空間．$(n = 1, 2, \ldots, N, N+1)$

A_n：決定空間で，Banach 空間．$(n = 1, 2, \ldots, N)$

$A_n(s_n) \subset A_n$：状態が $s_n \in S_n$ の時の可能決定空間．$(n = 1, 2, \ldots, N)$

$r_n : S_n \times A_n \to \mathbf{R}_+ \cup \{+\infty\} (\equiv \overline{\mathbf{R}}_+)$ 損失関数．$(n = 1, 2, \ldots, N)$

$k : S_{N+1} \to \overline{\mathbf{R}}_+$ 終端損失関数．

$T_n : S_n \times A_n \to 2^{S_{n+1}}$ 許容推移状態集合値写像．$(n = 1, 2, \ldots, N)$

初期状態 $s_1 \in S_1$ に対して，次のような最適化問題を考える．

$$(P) \quad \inf_{a_1} \mathop{\mathrm{Opt}}_{s_2} \inf_{a_2} \mathop{\mathrm{Opt}}_{s_3} \cdots \inf_{a_N} \mathop{\mathrm{Opt}}_{s_{N+1}} \sum_{n=1}^{N} r_n(s_n, a_n) + k(s_{N+1})$$
$$\text{subject to} \quad \begin{cases} a_n \in A_n(s_n), & n = 1, 2, \ldots, N \\ s_{n+1} \in T_n(s_n, a_n), & n = 1, 2, \ldots, N \end{cases}$$

ただし，Opt は，ここでは inf または sup を表している．

求めるべきものは，初期状態 $s_1 \in S_1$ に対する問題 (P) の最適値 $v_N(s_1)$ と，$v_N(s_1) = \sum_{n=1}^{N} r_n(\overline{s_n}, \overline{a_n}) + k(\overline{s_{N+1}})$ となる行動 $\overline{a_n} \in A_n(\overline{s_n})$ $(n = 1, 2, \ldots, N)$ と，状態 $\overline{s_{n+1}} \in T_n(\overline{s_n}, \overline{a_n})$ $(n = 1, 2, \ldots, N-1)$ である．

*The author is very grateful to Professor **K.Tanaka** of Niigata University and Professor **T.Tanaka** of Hirosaki University for their useful suggestions and encouragement on this research.

ここで，$n = 1, 2, \ldots, N+1$ に対して，$f_{N-n+1} : S_n \to \overline{\mathbf{R}}_+$を，

$$\begin{aligned} f_{N-n+1}(s_n) &= \inf_{a_n \in A_n(s_n)} \left\{ r_n(s_n, a_n) + \operatorname*{Opt}_{s_{n+1} \in T_n(s_n, a_n)} f_{N-n}(s_{n+1}) \right\}, \quad s_n \in S_n, \quad n = 1, 2, \ldots, N \\ f_0(s_{N+1}) &= k(s_{N+1}), \quad s_{N+1} \in S_{N+1} \end{aligned}$$

とすると，$f_N(s_1) = v_N(s_1)$, $\forall s_1 \in S_1$を満たす．よってf_{N-n+1}を，$(N-n+1)$段の最適値関数と呼ぶ．

このモデルでは，第n期において，状態s_nで決定a_nを選ぶと，次の状況s_{n+1}は集合$T_n(s_n, a_n)$の中から決まることになる．

この決まり方で代表的なものは2つあって，それは，決定者を最も有利にするものと，決定者を最も不利にするものであろう．前者は，例えば，決定者が状態を選べるときであり (最適子 Opt = inf)，後者は，例えば，ゲームの問題になる (最適子 Opt = sup)．

本報告では，最適子 Opt が特に infの場合について考察する．

2 最適子 Opt が infの時の解法

ここでは，最適子 Opt が infの時の，最適値関数列と最適行動，最適状態の列を考察する．

$(N-n+1)$段の最適値関数の関係式を簡略化して，次のように書く．

$$\varphi(f)(x) \equiv \inf_{y \in A(x)} \left\{ r(x,y) + \beta \inf_{z \in T(x,y)} f(z) \right\}, \; x \in X, \; f \in \overline{\mathbf{R}}_+^Z$$

ただし，$S_n = X$, $s_n = x$, $S_{n+1} = Z$, $s_{n+1} = z$, $A_n = Y$, $A_n(s_n) = A(x)$, $a_n = y$, $T_n = T$, $f_{N-n} = f$, $f_{N-n+1} = \varphi(f)$, $\beta \in (0, 1]$．

また，$\varphi^* : \overline{\mathbf{R}}_+^Z \to \overline{\mathbf{R}}_+^X$ を次のように定義する。$f \in \overline{\mathbf{R}}_+^Z$ に対して，

$$\varphi^*(f)(x) \equiv \inf_{(y^*, z^*) \in [\mathrm{Gr}T(x,\cdot)]^+} \{[r(x,\cdot) + \delta_{A(x)}]^*(y^*) + (\beta f)^*(z^*)\}$$

ただし，$\delta_{A(x)}(y) = \begin{cases} 0 & y \in A(x)\,; \\ +\infty & y \notin A(x)\,, \end{cases}$

$[\mathrm{Gr}T(x,\cdot)]^+ = \{(y^*, z^*) \in Y^* \times Z^* | \langle y, y^* \rangle + \langle z, z^* \rangle \geq 0, \; \forall (y,z) \in \mathrm{Gr}T(x,\cdot)\}$,

$[r(x,\cdot) + \delta_{A(x)}]^*(y^*) = \sup_{y \in Y} \{\langle y, y^* \rangle - [r(x,y) + \delta_{A(x)}(y)]\} = \sup_{y \in A(x)} \{\langle y, y^* \rangle - r(x,y)\}$,

$(\beta f)^*(z^*) = \sup_{z \in Z} \{\langle z, z^* \rangle - \beta f(z)\} = \beta \sup_{z \in Z} \left\{ \left\langle z, \frac{z^*}{\beta} \right\rangle - f(z) \right\} = \beta f^* \left(\frac{z^*}{\beta} \right)$ である．

Proposition 2.1 $\varphi(f)(x) < +\infty$ のとき，

$$\varphi(f)(x) + \varphi^*(f)(x) \geq 0$$

Proof. 任意の $y \in A(x)$, $z \in T(x,y)$, $(y^*, z^*) \in [\mathrm{Gr}T(x,\cdot)]^+$に対して，

$$\begin{aligned} 0 &\leq r(x,y) + \beta f(z) + \langle (y,z), (y^*, z^*) \rangle - r(x,y) - \beta f(z) \\ &= r(x,y) + \beta f(z) + \langle y, y^* \rangle - r(x,y) - \delta_{A(x)} + \langle z, z^* \rangle - \beta f(z) \\ &\leq r(x,y) + \beta f(z) + [r(x,\cdot) + \delta_{A(x)}]^*(y^*) + [\beta f(\cdot)]^*(z^*) \end{aligned}$$

上の不等式において，$y \in A(x)$, $z \in T(x,y)$, $(y^*, z^*) \in [\mathrm{Gr}T(x,\cdot)]^+$ に対する下限をとることにより，命題は示される． □

Theorem 2.1

$$(\theta_Y, \theta_Z) \in \mathrm{int}\mathrm{Gr}T(x,\cdot) - [\mathrm{dom}\, r(x,\cdot) \cap A(x)] \times \mathrm{dom} f$$

または,

$$(\theta_Y, \theta_Z) \in \mathrm{Gr}T(x,\cdot) - \mathrm{int}\{[\mathrm{dom}\, r(x,\cdot) \cap A(x)] \times \mathrm{dom} f\}$$

ならば,

$$\varphi^*(f)(x) = [r(x,\cdot) + \delta_{A(x)}]^*(\overline{y^*}) + (\beta f)^*(\overline{z^*})$$

となる $(\overline{y^*}, \overline{z^*}) \in [\mathrm{Gr}T(x,\cdot)]^+$ が存在する.

Proof. 仮定より, $\varphi(f)(x) < \infty$. これと, **Proposition 3.1**より, $\varphi^*(f)(x) > -\infty$である. もし, $\varphi^*(f)(x) = +\infty$ ならば, 任意の $(y^*, z^*) \in [\mathrm{Gr}T(x,\cdot)]^+$に対して $[r(x,\cdot) + \delta_{A(x)}]^*(y^*) + (\beta f)^*(z^*) = +\infty$ であるから, $\varphi^*(f)(x) < +\infty$ の時を示せばよい. このとき, $\varphi^*(f)(x)$ の定義より, 任意の自然数 $n \in \mathbf{N}$に対して,

$$[r(x,\cdot) + \delta_{A(x)}]^*(y_n^*) + (\beta f)^*(z_n^*) \leq \varphi^*(f)(x) + \frac{1}{n}$$

となる $(y_n^*, z_n^*) \in [\mathrm{Gr}T(x,\cdot)]^+$ が存在する. まず, この $\{(y_n^*, z_n^*)\}_{n=1}^{\infty}$ が有界であることを示す.

仮定より,

$$\delta B_Y \times \delta B_Z \subset \mathrm{Gr}T(x,\cdot) - [\mathrm{dom}\, r(x,\cdot) \cap A(x)] \times \mathrm{dom} f$$

となる正数 $\delta > 0$ が存在する. よって, 任意の $(u,v) \in B_Y \times B_Z$に対して, $\delta u = y_1 - y_2$, $\delta v = z_1 - z_2$ となる $(y_1, z_1) \in [\mathrm{dom}\, r(x,\cdot) \cap A(x)] \times \mathrm{dom} f$, $(y_2, z_2) \in \mathrm{Gr}T(x,\cdot)$ が存在する. ここで, 任意の $n \in \mathbf{N}$に対して,

$$\begin{aligned}
\delta\langle (u,v), (y_n^*, z_n^*)\rangle &= \langle y_1 - y_2, y_n^*\rangle + \langle z_1 - z_2, z_n^*\rangle \\
&= \langle y_1, y_n^*\rangle + \langle z_1, z_n^*\rangle - \langle y_2, y_n^*\rangle - \langle z_2, z_n^*\rangle \\
&= \langle y_1, y_n^*\rangle - [r(x,\cdot) + \delta_{A(x)}](y_1) + \langle z_1, z_n^*\rangle - \beta f(z_1) - \langle y_2, y_n^*\rangle - \langle z_2, z_n^*\rangle \\
&\quad + [r(x,\cdot) + \delta_{A(x)}](y_1) + \beta f(z_1) \\
&\leq [r(x,\cdot) + \delta_{A(x)}]^*(y_n^*) + [\beta f]^*(z_n^*) - \langle (y_2, z_2), (y_n^*, z_n^*)\rangle \\
&\quad + [r(x,\cdot) + \delta_{A(x)}](y_1) + \beta f(z_1) \\
&\leq [r(x,\cdot) + \delta_{A(x)}]^*(y_n^*) + [\beta f]^*(z_n^*) + [r(x,\cdot) + \delta_{A(x)}](y_1) + \beta f(z_1)
\end{aligned}$$

$\{[r(x,\cdot) + \delta_{A(x)}]^*(y_n^*) + (\beta f)^*(z_n^*)\}_{n=1}^{\infty}$ は, 有限値 $\varphi^*(f)(x)$ に収束する収束列であるから, 上式の右辺は n に関して有界である. よって, 一様有界性の定理より, $\{(y_n^*, z_n^*)\}_{n=1}^{\infty}$ は有界である.

Alaoglu の定理より, $\{(y_n^*, z_n^*)\}_{n=1}^{\infty}$は相対汎弱コンパクトであるから, $\{(y_n^*, z_n^*)\}_{n=1}^{\infty}$の部分列で, 汎弱収束するものがとれる. その部分列を $\{(y_{n'}^*, z_{n'}^*)\}$, 収束先を $(\overline{y^*}, \overline{z^*}) \in Y^* \times Z^*$とすると, $[\mathrm{Gr}T(x,\cdot)]^+$ は $Y^* \times Z^*$ の汎弱閉集合であるから, $(\overline{y^*}, \overline{z^*}) \in [\mathrm{Gr}T(x,\cdot)]^+$ である.

ここで, $[r(x,\cdot) + \delta_{A(x)}]^*$, $(\beta f)^*$ はそれぞれ Y^*, Z^*上で汎弱下半連続であるので,

$$\begin{aligned}
[r(x,\cdot) + \delta_{A(x)}]^*(\overline{y^*}) + (\beta f)^*(\overline{z^*}) &\leq \liminf_{n'\to\infty} \{[r(x,\cdot) + \delta_{A(x)}]^*(y_{n'}^*) + (\beta f)^*(z_{n'}^*)\} \\
&\leq \liminf_{n'\to\infty} \left\{\varphi^*(f)(x) + \frac{1}{n'}\right\} \\
&= \varphi^*(f)(x)
\end{aligned}$$

逆に, $\varphi^*(f)(x) \leq [r(x,\cdot) + \delta_{A(x)}]^*(\overline{y^*}) + (\beta f)^*(\overline{z^*})$ は明らかであるから, よって定理は示された. □

次に，$\varphi(f)(x)+\varphi^*(f)(x)=0$ となる十分条件を考える．一般に，$r(x,\cdot)\in\Gamma_0(Y)$, $f\in\Gamma_0(Z)$, $A(x)$ が closed convex set，$\mathrm{Gr}T(x,\cdot)$ が closed convex cone であり，かつ，**Theorem 3.1**の仮定が成立するならば，$\varphi(f)(x)+\varphi^*(f)(x)=0$ が成立する．(*cf.* J.-P.Aubin[2]) それよりも，もっと弱い十分条件を与える．そのため，$\Phi:\mathrm{dom}[r(x,\cdot)+\delta_{A(x)}]^*\times\mathrm{dom}(\beta f)^*\to\mathbf{R}\times 2^{Z^*}$ を次のように定義する．

$$\Phi(y^*,z^*)\equiv\left([r(x,\cdot)+\delta_{A(x)}]^*(y^*)+(\beta f)^*(z^*),\ K(y^*)-z^*\right)$$

ただし，$K(y^*)=\{z^*\in Z^*|(y^*,z^*)\in[\mathrm{Gr}T(x,\cdot)]^+\}$ とする．

Theorem 2.2 もし，$(-\varphi(f)(x),\theta_{Z^*})\in\mathrm{Im}\Phi+\left(\mathbf{R}_+\times\{\theta_{Z^*}\}\right)$ ならば，

$$\varphi(f)(x)+\varphi^*(f)(x)=0$$

かつ，

$$\varphi^*(f)(x)=[r(x,\cdot)+\delta_{A(x)}]^*(\overline{y^*})+(\beta f)^*(\overline{z^*})$$

となる $(\overline{y^*},\overline{z^*})\in[\mathrm{Gr}T(x,\cdot)]^+$ が存在する．

Proof. $(-\varphi(f)(x),\theta_{Z^*})\in\mathrm{Im}\Phi+\left(\mathbf{R}_+\times\{\theta_{Z^*}\}\right)$ とすると，$(-\varphi(f)(x),\theta_{Z^*})\in\left([r(x,\cdot)+\delta_{A(x)}]^*(\overline{y^*})+(\beta f)^*(\overline{z^*}),K(\overline{y^*})-\overline{z^*}\right)+(\overline{r},\theta_{Z^*})$ となる $(\overline{y^*},\overline{z^*})\in\mathrm{dom}[r(x,\cdot)+\delta_{A(x)}]^*\times\mathrm{dom}(\beta f)^*$ が存在する．これより，$-\varphi(f)(x)=[r(x,\cdot)+\delta_{A(x)}]^*(\overline{y^*})+(\beta f)^*(\overline{z^*})$, $(\overline{y^*},\overline{z^*})\in[\mathrm{Gr}T(x,\cdot)]^+$．従って，

$$\begin{aligned}0&=\varphi(f)(x)+[r(x,\cdot)+\delta_{A(x)}]^*(\overline{y^*})+(\beta f)^*(\overline{z^*})+\overline{r}\\&\geq\varphi(f)(x)+\varphi^*(f)(x)\geq 0\end{aligned}$$

よって，定理は示された． □

これより，次が示される．

Corollary 2.1 第 n 期において **Theorem 3.2**の仮定が成立しているならば，

$$f_{N-n+1}(s_n)=-\left\{[r_n(s_n,\cdot)+\delta_{A_n(s_n)}]^*(\overline{y_n^*})+[\beta f_{N-n}]^*(\overline{z_n^*})\right\}$$

を満たす $(\overline{y_n^*},\overline{z_n^*})\in[\mathrm{Gr}T_n(s_n,\cdot)]^+$ が存在する．

さらに，$f_{N-n+1}(s_n)=r_n(s_n,\overline{a_n})+\beta f_{N-n}(\overline{s_{n+1}})$ を満たす $\overline{a_n}\in A_n(s_n),\overline{s_{n+1}}\in T_n(s_n,\overline{a_n})$ が存在するならば，

$$\left\langle(\overline{a_n},\overline{s_{n+1}}),(\overline{y_n^*},\overline{z_n^*})\right\rangle=\left\langle\overline{a_n},\overline{y_n^*}\right\rangle+\left\langle\overline{s_{n+1}},\overline{z_n^*}\right\rangle=0$$

が成立する．

Proof. 系の前半は，**Theorem 3.2**より明らかである．

$\varphi(f)(x)=r(x,\overline{y})+\beta f(\overline{z})$, $\varphi^*(f)(x)=[r(x,\cdot)+\delta_{A(x)}]^*(\overline{y^*})+(\beta f)^*(\overline{z^*})$
$(\overline{y},\overline{z})\in\mathrm{Gr}T(x,\cdot)$, $(\overline{y^*},\overline{z^*})\in[\mathrm{Gr}T(x,\cdot)]^+$ とすると，

$$\begin{aligned}0&=\varphi(f)(x)+\varphi^*(f)(x)\\&=r(x,\overline{y})+\beta f(\overline{z})+[r(x,\cdot)+\delta_{A(x)}]^*(\overline{y^*})+(\beta f)^*(\overline{z^*})\\&\geq r(x,\overline{y})+\beta f(\overline{z})+\left\langle\overline{y},\overline{y^*}\right\rangle-r(x,\overline{y})+\left\langle\overline{z},\overline{z^*}\right\rangle-\beta f(\overline{z})\\&=\left\langle\overline{y},\overline{y^*}\right\rangle+\left\langle\overline{z},\overline{z^*}\right\rangle\\&=\left\langle(\overline{y},\overline{z}),(\overline{y^*},\overline{z^*})\right\rangle\\&\geq 0\end{aligned}$$

よって，示された. □

次に，段を可算無限個にしたものを考える．このときの最適化問題 (Q) は，初期状態 $s_1 \in X$ に対して，次のように与える.

$$(Q) \quad \inf_{a_1, s_2, a_2, s_3, \ldots} \sum_{n=1}^{\infty} \beta^{n-1} r(s_n, a_n)$$
$$\text{subject to} \quad \begin{cases} a_n \in A(s_n), & n = 1, 2, \ldots \\ s_{n+1} \in T(s_n, a_n), & n = 1, 2, \ldots \end{cases}$$

ただし,

X : 状態空間で，Banach 空間.

Y : 決定空間で，Banach 空間.

$A : X \to 2^Y$ 可能決定集合値写像.

$r : X \times Y \to \overline{\mathbf{R}}_+$ 損失関数.

$T : X \times Y \to 2^X$ 許容推移状態集合値写像.

β : 割引因子 $(0 < \beta < 1)$

とする．この問題 (Q) に対する，初期状態 $s_1 \in X$ のときの最適値を $v_\infty(s_1)$ とする.

Theorem 2.3 任意の $n = 1, 2, \ldots$ に対して，$\varphi^n(0)$ は，**Theorem 3.2**の仮定を満たすとする.

このとき，$\{\overline{y_n^*}\}_{n=1}^{\infty}, \{\overline{z_n^*}\}_{n=1}^{\infty}$ が存在して,

$$v_\infty(s_1) = \lim_{n \to \infty} f_n(s_1)$$

が成立する．ただし，$f_n \in \overline{\mathbf{R}}_+^X$ は次のように逐次的に作成される関数である.

$$f_0 \equiv 0, \ f_{n+1}(x) \equiv -\left\{[r(x, \cdot) + \delta_{A(x)}]^*(\overline{y_n^*}) + [\beta f_n]^*(\overline{z_n^*})\right\}$$

Proof. 明らかに，$\lim_{n \to \infty} f_n(x) = v_\infty(x)$ が成立する．これに，**Theorem 3.2**を適用する事により，定理は示される. □

また，$B(X) \equiv \{f \in \mathbf{R}^X \,|\, f$ は X の有界集合上で有界$\}$ とするとき，次がいえる.

Theorem 2.4 $\varphi : B(X) \to B(X)$ かつ，任意の $x \in X$ に対して,

$$\|z\| \leq \|x\|, \ \forall z \in T(x, y), \ \forall y \in A(x)$$

が成立しているとする．さらに $f_0 \equiv 0, \ f_{n+1} \equiv \varphi(f_n)$ とすると,

$$\lim_{n \to \infty} \|v_\infty - g_n\|_\infty = 0$$

が成立する．ただし，$f, g \in B(X)$ に対して,

$$\|f - g\|_k \equiv \sup_{x \in kB} |f(x) - g(x)|,$$

$$\|f - g\|_\infty \equiv \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{\|f - g\|_k}{1 + \|f - g\|_k}$$

である.

Proof. まず，$(B(X), \|\cdot\|_\infty)$ が Banach 空間になることは，よく知られている.

ここで，$x \in kB$とすると，

$$|\varphi(f)(x) - \varphi(g)(x)|_k \leq \beta \|f - g\|_k$$

であることが，比較的容易に確かめられる．これより，

$$\|\varphi(f) - \varphi(g)\|_\infty \leq \beta \|f - g\|_\infty$$

となる．$0 < \beta < 1$であるから，Banach-Picardの定理により，$\varphi(f) = f$を満たす$f \in B(X)$が存在する．ここで，$f_0 \equiv 0$，$f_{n+1} \equiv \varphi(f_n)$とするとき，

$$\begin{aligned} \|f_{n+1} - f\|_\infty &= \|\varphi(f_n) - \varphi(f\|_\infty) \\ &\leq \beta \|f_n - f\|_\infty \\ &\leq \cdots \leq \beta^{n+1} \|f_0 - f\|_\infty \end{aligned}$$

よって，

$$\lim_{n \to \infty} \|f_n - f\|_\infty = 0$$

ゆえに，

$$\lim_{n \to \infty} f_n(x) = f(x)$$

□

参考文献

[1] H. Attouch and H. Riahi, *Stability results for Ekeland's ϵ-variational principle and cone extremal solutions*, Math. Oper. Res. **18** (1993) 173–201.

[2] J- P. Aubin, *Optima and Equilibria* (Springer-Verlag, 1993).

[3] P. C. Bhakta and S. Mitra, *Some Existence Theorems for Functional Equations Arising in Dynamic Programming*, J. Math. Anal. Appl. **98** (1984) 348–362.

[4] P. C. Bhakta and S. R. Choudhury, *Some Existence Theorems for Functional Equations Arising in Dynamic Programming, II*, J. Math. Anal. Appl. **131** (1988) 217–231.

[5] P. Kanniappan, *Fenchel-Rockafellar Type Duality for a Non-Convex Non-Differential Optimization Problem*, J. Math. Anal. Appl. **97** (1983) 266–276.

[6] 田中謙輔 et al., *Fenchel Duality* の応用, to appear in 数理科学講究録 (1994).

[7] R. T. Rockafellar, *Extension of Fenchel's Duality Theorem for Convex Functions*, Duke Math. J. **33** (1966) 81–89.